

**Methodology for the Development
of the
2006 Section 303(d) List in Missouri**

Missouri Department of Natural Resources
Division of Environmental Quality
Water Protection Program

Approved by the Clean Water Commission June 7, 2006

Table of Contents

I.	Citation and Requirements	2
	A. Citation of Section of Clean Water Act.....	2
	B. USEPA Guidance..... Placement of Waters Within the Five 305(b) Report Categories.....	2
II.	Methodology Document	5
	A. Procedures and Methods Used to Collect Water Quality Data.....	5
	B. Identification of All Existing and Readily Available Water Quality Data Sources.....	9
	C. Data Quality Considerations	11
	D. How Water Quality Data is Evaluated to Determine Whether or Not Waters are Impaired.....	13
	E. 303(d) Listing Considerations.....	17
	Appendix A: Excerpts from EPA Guidance on Statistical Approaches.....	18
	Appendix B: Statistical Considerations.....	20
	Appendix C: Examples of Statistical Procedures.....	25

Tables

Table 1: Methods for Assessing Compliance with Water Quality Standards.....	13
Table B-1: Description of Analytical Tools	21
Table B-2: Effects of Type 1 Error Rates and Sample Size on Type 2 Error Rates.....	24

I. Citation and Requirements

A. Citation of Section of Clean Water Act

This document is required by revisions of rules under the federal Clean Water Act, Section 303(d), 40 CFR 130.7 and the timetable for presenting the finished document to EPA and the public is given in part 130.10. Section 303(d) requires states to list certain impaired waters and the rules require that states describe how this list will be constructed.

B. USEPA Guidance

In July 2003, USEPA issued new guidance, the “Guidance for 2004 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d) and 305(b) of the Clean Water Act (CWA).” This guidance gives further recommendations about listing of 303(d) and other waters. In July 2005, USEPA published an amended version, “Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act.”

The entity responsible for administration of the federal Clean Water Act in Missouri is the Missouri Department of Natural Resources (the department). EPA regulations require that the department describe the methodology used to develop the state’s 303(d) list. This draft document should be made available by the department to the public for at least a 60-day public review and comment period. The department should provide the USEPA with a document summarizing all comments received and department “responses to significant comments.” The EPA guidance recommends that the department provide “a description of the methodology used to develop the section 303(d) list, (2) a description on the data and information used to identify (impaired and threatened) waters, including a description of the existing and readily available data and information used, and (3) a rationale for any decision to not use any existing and readily available data and information. The guidance also notes that “prior to submission of its Integrated Report, each state should provide the public with the opportunity to review and comment on the methodology”. The guidelines further recommend that the Methodology Document include information on how interstate/international disagreements concerning the list are resolved.

Placement of Waters within the Five Categories in the 2006 EPA Assessment, Listing and Reporting guidance.

The guidance issued by EPA in 2005 recommends that all the classified waters of the state be placed in one of five categories.

Category 1

All designated beneficial uses are fully maintained. Data or other information supporting full beneficial use attainment for all designated beneficial uses must be consistent with the state listing methodology document. The department will place a water in Category One if both of the following conditions are met:

- The water has chemical (at a minimum, water temperature, pH, dissolved oxygen and ammonia) and/or biological water quality data (at a minimum, E. coli or fecal coliform bacteria) that indicate attainment with all designated uses, or the watershed of this water has geology and land use similar to a water that does have water quality data that indicate attainment with all designated uses
- The level of mercury in fish fillets does not exceed fish tissue guidelines of 0.3 mg/kg or less. Where at least three samples are available for higher trophic level species (bass, sauger, walleye, northern pike), only those samples will be used. When there are less than three samples from these species, data from all fish species will be used.

Category 2

One or more designated beneficial uses are fully attained but at least one designated beneficial use has inadequate data or information to make a use attainment decision consistent with the state listing methodology document. The department will place a water in Category Two if it cannot be placed in Category One due to the lack of a comparable watershed with similar geology and land use in Category One, and at least one of the following conditions are met:

- There is inadequate data for water temperature, pH, dissolved oxygen or ammonia to assess attainment with the aquatic life protection use.
- There is inadequate e coli or fecal coliform bacteria data to assess attainment with the whole body contact recreational use. There is insufficient fish fillet tissue data available for mercury to assess attainment with the fish consumption use.

Category 3

None of the designated beneficial uses have adequate data or information to make a use attainment decision consistent with the state listing methodology document.

Category 3A. The department will place a water in Category 3A if it cannot be placed in Category One due to the lack of a comparable watershed with similar geology and land use in Category One, and if either of the following conditions are met:

- The water is currently not assessed for compliance with state water quality standards due to the absence of water quality data for this water.
- There is some data available for the water, but it is not sufficient to perform an assessment, and it does not suggest the presence of any impairment.

Category 3B. The department will place a water in Category 3B if the following conditions are met:

- Available data suggests potential noncompliance with state water quality standards but more data is needed to meet listing methodology document data requirements or to make conclusive use attainment decisions. This would include inconclusive statistical tests for small data sets or Level One, Two, Three or Four data not supported by adequate quality assurance information.

Category 4

State water quality standards as per the requirements of Table One are not attained, but a TMDL is not required.

Category 4A

EPA has approved a TMDL that addresses the impairment. The department will place a water in Category 4A if both the following conditions are met:

- Any portion of the water is rated as being in non-attainment with state water quality standards as per the requirements of Table One due to one or more discrete pollutants or discrete properties of water¹, and
- USEPA has approved a TMDL for all pollutants causing that non-attainment.

Category 4B

Water pollution controls required by a local, state or federal authority, are expected to correct the impairment in a reasonable period of time. The department will place a water in Category 4B if both of the following conditions are met:

- Any portion of the water is rated as being in non-attainment with state water quality standards as per the requirements of Table One due to one or more discrete pollutants or discrete properties of water, and
- A water quality based permit that addresses the pollutant(s) causing the designated use impairment has been issued or is expected to be issued within the assessment cycle. Or other pollution control requirements have been made or will be made during the assessment cycle that are expected to adequately address the pollutant(s) causing the impairment. This may include implemented voluntary watershed control plans as enunciated in the EPA guidance document.

Category 4C

A discrete pollutant(s) or other discrete property of the water does not cause the impairment. The department will place a water in Category 4C if the following conditions are met:

- Any portion of the water is rated as being in non-attainment with state water quality standards as per the requirements of Table One, and
- The only cause(s) of the non-attainment is not a specific chemical element (E.g., lead, zinc), chemical compound (E.g., ammonia, dieldrin, atrazine) or one of the following quantifiable physical, biological or bacteriological conditions: water temperature, percent of gas saturation, amount of dissolved oxygen, pH, deposited sediment, suspended sediment, turbidity, toxicity or density of fecal coliform or E. coli bacteria.
- If a designated use is not supported and the segment is impaired or threatened, the fact that a specific pollutant is not known does not provide a basis for excluding a segment from Category 5. These segments must be listed unless the state can

¹ A discrete pollutant or a discrete property of water is defined here as a specific chemical or other attribute of the water (such as temperature, dissolved oxygen or pH) that causes beneficial use impairment and that can be measured quantitatively.

demonstrate that no pollutant or pollutants causes or contributes to the impairment or to establishing a TMDL for such segments, the pollutant causing the impairment must be identified.

Category 5

At least one discrete pollutant has caused non-attainment with state water quality standards as per the requirements of Table One and the water does not meet the criteria for either category 4A or 4B. Category 5 waters are the state 303(d) listed waters. The department will place a water in Category 5 if both the following conditions are met:

- Any portion of the water is rated as being in non-attainment with state water quality standards as per Table One due to one or more discrete pollutants or discrete properties of the water, and
- That portion of the water is not listed in Category 4A or 4B for that pollutant.

Threatened Waters

When a water that would otherwise be in Category One, Two or Three has a time trend analysis for one or more discrete water quality pollutants that indicates the water is currently maintaining all beneficial uses but will not continue to meet these uses before the next listing cycle, it will be considered a “threatened water”. A threatened water will be treated as an impaired water and placed in the appropriate category (4A,4B or 5).

II. The Methodology Document

A. Procedures and Methods Used to Collect Water Quality Data

MDNR Monitoring

The major purposes of the Department of Natural Resources' water quality monitoring program are (1) to characterize background or reference water quality conditions; (2) to better understand daily, flow event, and seasonal water quality variations and their underlying processes; (3) to characterize aquatic biological communities (4) to assess time trends in water quality; (5) to characterize local and regional impacts of point and nonpoint source discharges on water quality; (6) to check for compliance with water quality standards or wastewater permit limits and; (7) to support development of strategies, including TMDLs, to return impaired waters to compliance with water quality standards. All of these objectives are statewide in scope.

Coordination with Other Monitoring Efforts in Missouri

The Department of Natural Resources cooperates with other agencies in performing special water quality studies.

In 1998, a multi-agency monitoring task force including the Missouri Department of Natural Resources, the Missouri Department of Conservation, the U.S. Environmental Protection Agency, the U.S. Geological Survey, the U.S. Forest Service, the USDA Natural Resources Conservation Service and the University of Missouri convened. The goals of this group were to

outline a statewide aquatic resources monitoring plan, define partnership roles in this monitoring plan and to discuss the kind of research needed to further this new monitoring effort. A review of the entire Water Pollution Control Program monitoring plan by USEPA began in June 2003. Several state and federal agencies will participate in this review.

To maximize efficiency, the department routinely coordinates its monitoring activities to avoid overlap with other agencies and provide and receive interagency input on monitoring study design. Data from other sources is used for meeting the same objectives as department sponsored monitoring. The agencies most often involved are the U.S. Geological Survey (USGS), the U.S. Army Corps of Engineers, the U.S. Environmental Protection Agency (EPA), the Missouri Department of Conservation (MDC), the USDA/Agricultural Research Service (ARS) and the Missouri Department of Health and Senior Services. However, the department also tracks the monitoring efforts of the National Park Service, the U.S. Forest Service, several of the state's larger cities, the states of Oklahoma, Arkansas, Kansas, Iowa and Illinois and graduate level research conducted at universities within Missouri. For those wastewater discharges where the department has required instream water quality monitoring, the department may also use monitoring data acquired by wastewater dischargers as a condition of discharge permits issued by the department. In 1995, the department began using data collected by volunteers that have passed Quality Assurance/Quality Control (QA/QC) tests.

Existing Monitoring Networks and Programs

The following is a list description of the kinds of water quality monitoring activities presently occurring in Missouri.

1. Fixed Station Network

- A. Objective: To better characterize background or reference water quality conditions, to better understand daily, flow event and seasonal water quality variations and their underlying processes, to assess time trends and to check for compliance with water quality standards.
- B. Design Methodology: Sites were chosen based on one of the following criteria:
 - site is believed to have water quality representative of many neighboring streams of similar size due to similarity in watershed geology, hydrology and land use, and the absence of any impact from a local point or discrete nonpoint water pollution source.
 - site is downstream of a significant point source or localized nonpoint source area.
- C. Number of Sites, Sampling Methods, Sampling Frequency, Parameters:
 - USGS/DNR cooperative network: 57 sites statewide, horizontally and vertically integrated grab samples, 6-12 times per year for major ions, nutrient ions, temperature, pH, dissolved oxygen, specific conductance and flow, 2-4 times annually for suspended solids and heavy metals, and for pesticides 6 times annually at 6 sites.
 - DNR raw water sampling of public drinking water reservoirs: nine drinking water reservoirs are sampled 4 times/year for some commonly used agricultural herbicides.
 - UMC/DNR lake monitoring network. This program has monitored about 110 lakes. About forty lakes are monitored each year. Each lake is usually sampled four times during the summer and about 12 are monitored spring through fall for nutrients,

- chlorophyll, turbidity and suspended solids.
- DNR routine monitoring of finished public drinking water supplies for bacteria and trace contaminants.
 - Routine bacterial monitoring (typically weekly during the summer) of swimming beaches at Missouri state parks during the recreational season by DNR, Division of State Parks.
 - Monitoring of sediment quality by DNR at approximately ten discretionary sites annually. All sites are monitored for several heavy metals and organic contaminants. A pore water sample is analyzed for ammonia and a Microtox toxicity test is performed on the sediment.

2. Special Water Quality Studies

- A. Objective: Special water quality studies are used to characterize the water quality impacts from a specific pollutant source area.
- B. Design Methodology: These studies are designed to determine the contaminants of concern based on previous water quality studies, effluent sampling and/or Missouri State Operating Permit applications. These studies employ multiple sampling stations downstream and upstream (if appropriate). If contaminants of concern have significant seasonal or daily variation, season of the year and time of day variation must be accounted for in the sampling design.
- C. Number of Sites, Sampling Methods, Sampling Frequency, Parameters: The Missouri Department of Natural Resources conducts or contracts for 10-15 special studies annually. Each study has multiple sampling sites. Number of sites, sampling frequency and parameters all vary greatly depending on the study. Intensive studies would also require multiple samples per site over a relatively short time frame.

3. Toxics Monitoring Program

The fixed station network and many of our intensive studies monitor for toxic chemicals. In addition, major municipal and industrial dischargers must monitor for toxicity in their effluents as a condition of their Missouri State Operating Permits.

4. Biological Monitoring Program

- A. Objectives: The objectives of this program are to develop numeric criteria describing “reference” aquatic macroinvertebrate and fish communities in Missouri’s streams, to implement these criteria within state water quality standards and to continue a statewide fish and aquatic invertebrate monitoring program.
- B. Design Methodology: Development of biocriteria for invertebrates and fish involves identification of streams divided among Missouri’s aquatic ecoregions. It also includes intensive sampling of invertebrate and fish communities to quantify temporal and spatial variation in reference streams within ecoregions and variation between ecoregions, and the sampling of chemically and physically impaired streams to test sensitivity of various community metrics to differences in stream quality.

C. Number of Sites, Sampling Methods, Sampling Frequency, Parameters:

The department has conducted biological sampling of aquatic invertebrates for many years. Since 1991 this program has consisted of standardized monitoring of approximately 55 sites twice annually. The Missouri Department of Conservation presently has a statewide fish and aquatic invertebrate monitoring program designed to assess and monitor the health of Missouri's stream resources. This program samples a minimum of 450 random and 30 reference sites every five years.

5. Fish Tissue

A. Objective: Fish tissue monitoring can address two separate objectives. These are 1) the assessment of ecological health or the health of aquatic biota (usually accomplished by monitoring whole fish samples) and 2) the assessment of human health risk based on the level of contamination of fish fillets.

B. Design Methodology. Fish tissue monitoring sites were chosen based on one of the following criteria:

- site is believed to have water and sediment quality representative of many neighboring streams of similar size due to similarity in geology, hydrology and land use, and the absence of any known impact from a local point source or discrete nonpoint water pollution source.
- site is downstream of a significant point source or localized nonpoint source area.

C. Number of Sites, Sampling Methods, Sampling Frequency, Parameters:

Missouri DNR and USEPA have a cooperative fish tissue monitoring program that collects whole fish composite samples² at approximately 12 fixed sites. Each site is sampled once every two years. The preferred species for these sites are either carp or redhorse sucker. This program also samples approximately five discretionary sites annually for two fish fillet composite samples. One sample is of a top carnivore fish such as largemouth bass, smallmouth bass, walleye or sauger. The other sample is for a species of a lower trophic order such as catfish, carp or sucker.

This program also samples forty to fifty discretionary sites annually for two fish fillet composite samples. One sample is of a top carnivore such as largemouth bass, smallmouth bass, walleye or sauger. The other sample is for a species of a lower trophic level such as catfish, carp or sucker. The Missouri Department of Conservation is a partner in this portion of the program.

The Missouri Department of Conservation samples approximately 31 discretionary sites annually for fish fillets. They sample a wide variety of species. Both of these monitoring programs analyze for several chlorinated hydrocarbon insecticides, PCBs, lead, cadmium, mercury and fat content.

6. Volunteer Monitoring Program

Two volunteer monitoring programs are now generating water quality data in Missouri. The first is a cooperative program between the Department of Natural Resources, the University

² A composite sample is one in which several individual fish are combined to produce one sample.

of Missouri and volunteers that monitors approximately 50 lakes, including Lake Taneycomo, Table Rock Lake and several lakes in the Kansas City area. Data from this program is used by the University as part of a long-term study on the limnology of midwestern reservoirs.

The second program involves volunteers who monitor water quality of streams throughout Missouri. The Volunteer Water Quality Monitoring Program is a cooperative project of the Department of Natural Resources, the Department of Conservation and the Conservation Federation of Missouri. By the end of 2003, just over 3000 citizen volunteers had attended at least one training workshop. After the introductory class, many proceed on to at least one more class of higher level training; Levels 1,2,3 and 4. Each level of training is a prerequisite for the next higher level, as is appropriate data submission. Training Levels 2, 3 and the newest Level, 4 (piloted in 2003) represent increasingly higher quality assurance. Of those completing an introductory course, 383 (about 13%) proceed to Levels 1 and 2. Seventy-two volunteers have reached Level 3.

Laboratory Analytical Support

Laboratories used:

- USGS/DNR Cooperative Fixed Station Network: USGS Lab, Denver, Colorado
- DNR Public Drinking Water Reservoir Network: Missouri DNR Environmental Lab
- Intensive Surveys: Varies, many are done by Missouri DNR Environmental Lab
- Toxicity Testing of Effluents: many commercial labs
- Biological Criteria for Aquatic Invertebrates: Missouri DNR Environmental Lab and University of Missouri, Columbia
- Fish Tissue: USEPA Region VII Lab, Kansas City, Kansas and Miscellaneous contract labs (Missouri Department of Conservation)
- Missouri State Operating Permit: self-monitoring or commercial labs
- DNR Public Drinking Water Monitoring: Missouri DNR and commercial labs
- Agricultural Research Service: ARS lab
- Other water quality studies: many commercial labs

B. Identification of All Existing and Readily Available Water Quality Data Sources

The following data sources are used by the department to aid in the compilation of the state 305(b) report. Where quality assurance programs are deemed acceptable, these sources would also be used to develop the state Section 303(d) list. These sources presently include but are not limited to:

1. Fixed station water quality and sediment data collected and analyzed by Missouri DNR personnel in the Missouri DNR Lab.
2. Fixed station water quality data collected by the U.S. Geological Survey under contractual agreements with Missouri DNR.
3. Fixed station water quality data collected by the U.S. Geological Survey under contractual agreements to agencies or organizations other than Missouri DNR.
4. Fixed station water quality, sediment quality and aquatic biological information collected by the U.S. Geological Survey under their NASQAN and NAWQA monitoring programs.
5. Fixed station raw water quality data collected by the Kansas City Water Services Department, the St. Louis City Water Company, St. Louis County Water Company,

Springfield City Utilities and Springfield Department of Public Works.

6. Fixed station water quality data collected by the U.S. Army Corps of Engineers. The St. Louis, Kansas City and Little Rock districts have monitoring programs for Corps-operated reservoirs in Missouri.
7. Fixed station water quality data collected by the Arkansas Department of Environmental Quality, the Kansas Department of Health and Environment and the Iowa Department of Natural Resources and the Illinois Environmental Protection Agency.
8. Fixed station water quality monitoring by corporations.
9. Annual fish tissue monitoring programs by USEPA/Missouri Department of Natural Resources (RAFTMP program) and the Missouri Department of Conservation.
10. Special Water Quality Surveys conducted by Missouri Department of Natural Resources. Most of these surveys are focused on the water quality impacts of specific point source wastewater discharges. Some surveys are of well-delimited nonpoint sources such as abandoned mined lands. These surveys often include physical habitat evaluation and monitoring of aquatic invertebrates as well as water chemistry monitoring.
11. Special Water Quality Surveys conducted by the U.S. Geological Survey, including but not limited to:
 - a) geological, hydrological and water quality of various hazardous waste sites
 - b) geological, hydrological and water quality of various abandoned mining areas
 - c) hydrology and water quality of urban nonpoint source runoff in St. Louis, Kansas City and Springfield, Missouri
 - d) Bacterial and nutrient contamination of streams in southern Missouri
12. Special water quality studies by other agencies such as the Missouri Department of Conservation, the U.S. Public Health Service, and the Missouri Department of Health and Senior Services.
13. Monitoring of fish occurrence and distribution by the Department of Conservation.
14. Fish Kill and Water Pollution Investigations Reports published by the Missouri Department of Conservation.
15. Selected graduate research projects pertaining to water quality and/or aquatic biology.
16. Water quality, sediment and aquatic biological data collected by the department, USEPA or their contractors at hazardous waste sites in Missouri.
17. Self-monitoring of receiving streams by cities, sewer districts and industries, or contractors on their behalf, for those discharges that require this kind of monitoring. This monitoring includes chemical and sometimes toxicity monitoring of some of the larger wastewater discharges, particularly those that discharge to smaller streams and have the greatest potential to affect instream water quality.
18. Compliance monitoring of receiving waters by the Missouri Department of Natural Resources and USEPA. This can include chemical and toxicity monitoring.
19. Bacterial monitoring of streams and lakes by county health departments, community lake associations and other organizations using acceptable analytical methods.
20. Other monitoring activities done under a quality assurance project plan approved by the department.

The following data sources cannot be used to rate a water as impaired (Category 4A, 4B, 4C or 5). However these data sources may be used to direct additional monitoring that would allow a water quality assessment for Section 303(d) listing purposes.

21. Fixed station water quality and aquatic invertebrate monitoring by volunteers who have successfully completed the Volunteer Monitoring Program Level Two

- workshops. Data collected by volunteers who have successfully completed a training Level 2 workshop is considered to be data quality code Level One. This data is eligible for use in assigning waters to categories 1,2,3A or 3B but not to categories 4A, 4B, 4C or 5. Data from volunteers who have not yet completed a Level 2 training workshop do not have sufficient quality assurance to be used for any assessment purposes.
22. Fish Management Basin Plans published by the Missouri Department of Conservation.
 23. Fish Consumption Advisories published annually by the Missouri Department of Health and Senior Services.
 24. Self-monitoring of wastewaters by cities, sewer districts and industries, or contractors on their behalf, that have significant wastewater discharges. This monitoring includes chemical and sometimes toxicity monitoring of some of the larger wastewater discharges, particularly those that discharge to smaller streams and have the greatest potential to affect instream water quality.
 25. Compliance monitoring of wastewaters by the Missouri Department of Natural Resources and USEPA. This can include chemical and toxicity monitoring.

C. Data Quality Considerations

Quality Assurance/Quality Control Programs

1. Missouri Department of Natural Resources Quality Assurance/Quality Control Program

Missouri DNR and USEPA Region VII have completed a Total Quality Management Plan. All environmental data generated directly by the department or through contracts funded by the department or EPA require a Quality Assurance Project Plan (QAPP). The agency or organization responsible for collection or collection and analysis of the environmental sampling must write and adhere to a QAPP approved through the Missouri Department of Natural Resources' Quality Management Plan.

2. Other Quality Assurance/Quality Control Programs

Water quality data from any source will be judged as "scientifically defensible" if the data was generated under a quality assurance project plan approved by the department. Data generated in the absence of a department-approved quality assurance project plan may be used to determine the 303(d) status of a water if the department determines that the data is scientifically defensible after making a review of the quality assurance procedures used by the data generator.

Other Data Quality Considerations

1. Data Age

For assessing present conditions, more recent data is preferable but older data can be used to assess present conditions if the data remains representative of present conditions.

If the department uses data to make a 303(d) listing decision that predates the date the list is developed by more than seven years, the department will provide a written justification for the use of such data.

A second consideration is the age of the data relative to significant events that may have an effect on water quality. Data collected prior to the initiation, closure or significant change in a wastewater discharge, or prior to a large spill event or the reclamation of a mining or hazardous waste site, for example, is not representative of present conditions. Such data would not be used to assess present conditions even if it was less than seven years old. Such “pre-event” data can be used to determine changes in water quality before and after the “event” or to show water quality time trends.

2. Data Type, Amount and Information Content

The USEPA recommends establishing a data code and rating data quality by the kind and amount of data present at a particular location (USEPA 1997³). The code is a single digit number from one to four, indicating the relative degree of assurance the user has in the value of a particular environmental data set. Level One indicates the least assurance and Level Four the greatest. Based on EPA’s guidance, the Missouri Department of Natural Resources uses the following rules to assign code numbers to data.

- Level One: - All data not meeting the requirements of Levels Two, Three or Four.
- Level Two: - Chemical data collected quarterly to bimonthly for at least three years or intensive studies that monitor several nearby sites repeatedly over short periods of time or
 - At least three fish tissue samples per waterbody.
- Level Three: - Chemical data collected at least monthly for more than three years and providing data on a variety of water quality constituents including heavy metals and pesticides; or
 - Quantitative biological monitoring of at least one aquatic assemblage (fish, invertebrates or algae) at multiple sites, or multiple samples at a single site when data from that site is supported by biological monitoring at an appropriate control site.
- Level Four: - Chemical data collected at least monthly for more than three years and providing data on a variety of water quality constituents including heavy metals and pesticides, and including chemical sampling of sediments and fish tissue; or
 - Quantitative biological monitoring of at least two aquatic assemblages (fish, invertebrates or algae) at multiple sites.

In Missouri, the primary purpose of Level One data is to provide a rapid and inexpensive method of screening large numbers of waters for obvious water quality problems and to determine where more intensive monitoring is needed. In the preparation of the state 305(b) report, data from all four data quality levels are used. Most of the data is of Level One quality, and without Level One data staff would not be able to assess a majority of the state’s waters.

³ Guidelines for the Preparation of the Comprehensive State Water Quality Assessments (305b) and Electronic Updates, 1997.

In general, when selecting waterbodies for the Missouri 303(d) list, only Level Two or higher data are used, unless the problem can be accurately characterized by Level One data.⁴ The reason is that Level Two data provides a higher level of assurance that a water quality standard is actually being exceeded and that a TMDL study is necessary. All waterbodies placed in Category 3B receive high priority for additional monitoring so that data quality is upgraded to at least Level Two.

D. How Water Quality Data is Evaluated to Determine Whether or Not Waters are Impaired for 303(d) Listing Purposes

Physical, Chemical, Biological and Toxicity Data

USEPA provides guidelines to the states on how to evaluate water quality data in order to determine if waters are impaired (USEPA 2005⁵). These guidelines, shown in Table One below, are those that are used to determine if individual waters are impaired. In addition, if time trend data indicates that presently unimpaired waters will become impaired prior to the next listing cycle, these “threatened waters” will be judged to be impaired. Where antidegradation provisions in Missouri’s water quality standards apply, those provisions shall be upheld.

Table 1. METHODS FOR ASSESSING COMPLIANCE WITH
WATER QUALITY STANDARDS USED FOR 303(d)LISTING PURPOSES

BENEFICIAL USES	DATA TYPE	DATA QUALITY CODE	COMPLIANCE WITH WATER QUALITY STANDARDS
Overall use protection (all beneficial uses)	No data--evaluated based on similar land use/geology as stream with water quality data. ⁶	na	Given same rating as monitored stream with same land use and geology.
Any beneficial uses.	No data available or where only effluent data is available. Results of dilution calculations or water quality modeling. (see ALRR p.38)	na	Where models or other dilution calculations indicate noncompliance with allowable pollutant levels and frequencies noted in this table, waters may be added to category 3B and considered high priority for water quality monitoring.
Overall use protection (all beneficial uses)	Narrative criteria for which quantifiable measurements can be made.	1	<u>Full:</u> Stream appearance typical of reference streams in this region of the state. <u>Non-Attainment:</u> Presence of objectionable or unsightly color, odor, turbidity, bottom deposits, oil, scum, floating or suspended debris, or the presence of substances in sufficient amounts to prevent full

⁴ When a listing, amendment or de-listing of a 303(d) water is made with only Level One data, a document will be prepared that includes a display of all data and a presentation of all statistical tests or other evaluative techniques that documents the scientific defensibility of the data. This requirement applies to all Level One data identified in Table One.

⁵ Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and Section 314 of the Clean Water Act. USEPA.

⁶ This data type is used only for wide-scale assessments of aquatic biota and aquatic habitat for 305(b) report purposes. This data type is not used in the development of the 303(d) list.

			<p>maintenance of beneficial uses. For the purposes of 303(d) listing; a waterbody will be considered to be not in conformance with narrative water quality criteria if these quantifiable limits are exceeded. Acute criteria for aquatic life shall not be exceeded more than one time in three years.</p> <p>Color: color as measured by the Platinum-Cobalt visual method (SM 2120 B) in a water is statistically significantly higher than a control water.</p> <p>Objectionable Bottom Deposits: the affected stream segment must be at least 100 yards in length, and for all areas within this affected segment that have a flow velocity less than 0.5 feet/second at the time the stream is evaluated, greater than 10% of the stream bottom is covered by sewage sludge, trash or other materials reaching the water due to anthropogenic sources.</p> <p>Note: Waters in mixing zones and unclassified waters which support aquatic life on an intermittent basis shall be subject to acute toxicity criteria for protection of aquatic life. Waters in the intial Zone of Dilution (ZID) shall not be subject to acute toxicity criteria.</p>
Protection of Aquatic Life	Toxic Chemicals	1-4	<p><u>Full:</u> No more than 1 acute toxic event in three years. No more than one exceedence of acute or chronic criterion in 3 years for all toxics other than ammonia; for ammonia , no more than one exceedence of acute criterion in 3 years and 10% or fewer of all samples exceed chronic criterion.^{7,8}</p> <p><u>Non-Attainment:</u> Requirements for full attainment not met. (see CALM p.27,30. ALRR p.39).</p>
Protection of Aquatic Life	Conventional (pH, temperature, dissolved oxygen, total dissolved gases, oil and grease, sulfate plus chloride)	1-4	<p><u>Full:</u> No more than 10% of all samples exceed criterion⁷.</p> <p><u>Non-Attainment:</u> Requirements for full attainment not met.</p>

⁷ The time period used to calculate the average will be the entire period that is judged to representative of present conditions. See section on Data Age.

BENEFICIAL USES	DATA TYPE	DATA QUALITY CODE	COMPLIANCE WITH WATER QUALITY STANDARDS
Protection of Aquatic Life	Biological	3-4	<p><u>Full:</u> Fauna very similar to regional reference streams. If DNR wadeable streams macroinvertebrate sampling and evaluation protocols are followed, for seven or fewer samples, at least 75% of the stream condition index scores must be 16 or greater. For greater than seven samples or for other sampling and evaluation protocols, results must be statistically similar to representative reference or control streams.</p> <p><u>Non-Attainment:</u> If DNR wadeable streams macroinvertebrate sampling and evaluation protocols are followed, for seven or fewer samples, at least 75% of the stream condition index scores must be 14 or lower. For more than seven samples or for other sampling and evaluation protocols, results must be statistically dissimilar to representative reference or control streams.</p>
Protection of Aquatic Life	Toxicity testing of streams or lakes using aquatic organisms.	2	<p><u>Full:</u> No more than one test result of statistically significant deviation from controls in acute or chronic test in a 3-year period. ⁸</p> <p><u>Non-Attainment:</u> Requirements for full attainment not met.</p>
Fish Consumption	Chemicals (water)	1-4	<p><u>Full:</u> Water quality does not exceed water quality standard.⁹</p> <p><u>Non-Attainment:</u> Requirements for full attainment not met.</p>
Fish Consumption	Chemicals (tissue)	1-2	<p><u>Full:</u> Fish tissue levels in fillets do not exceed guidelines¹⁰</p> <p><u>Non-Attainment:</u> Requirements for full attainment not met.</p>

⁸ The test result must be representative of water quality for the entire time period for which acute or chronic criteria apply. The department will review all appropriate data, including hydrographic data, to insure only representative data is used. Except on large rivers where stormwater flows may persist at relatively unvarying levels for several days, grab samples collected during stormwater flows will not be used for assessing chronic toxicity criteria.

⁹ See section on Statistical Considerations and Table B-1.

¹⁰ Fish tissue threshold levels are Chlordane 0.1 mg/kg (Crellin, J.R. 1989, "New Trigger Levels for Chlordane in Fish-Revised Memo" Mo. Dept. of Health interoffice memorandum. June 16, 1989), Mercury 0.3 mg/kg (based on two documents: Mercury Levels in Commercial Fish and Shellfish- <http://www.cfsan.fda.gov/~frf/sea-mehg.html> and FDA and EPA Announce the Revised Consumer Advisory on Methylmercury in Fish- <http://www.fda.gov/bbs/topics/news/2004/NEW01038.html>), PCBs 2.0 mg/kg (USFDA Industries Activities Staff Booklet, August 2000 (<http://vm.cfsan.fda.gov/~lrd/daact.html>) , and Lead 0.3 mg/kg (World Health Organization 1972. "Evaluation of Certain Food Additives and the Contaminants Mercury, Lead and Cadmium". WHO Technical

BENEFICIAL USES	DATA TYPE	DATA QUALITY CODE	COMPLIANCE WITH WATER QUALITY STANDARDS
Drinking Water Supply ¹¹ -Raw Water.	Chemical (toxics)	1-4	<u>Full</u> : Water Quality Standard not exceeded. ⁹ <u>Non-Attainment</u> : Requirements for full attainment not met.
Drinking Water Supply- Raw Water ¹¹	Chemical (Total Dissolved Solids)	1-4	<u>Full</u> : Water Quality Standard not exceeded. ⁹ <u>Non-Attainment</u> : Requirements for full attainment not met
Drinking Water Supply-Finished Water	Chemical (toxics)	1-4	<u>Full</u> : No MCL* violations based on SDWA data evaluation procedures. <u>Non-Attainment</u> : Requirements for full attainment not met. NOTE: Finished water data will not be used for analytes where water quality problems may be caused by the drinking water treatment process such as the formation of Trihalomethanes (THMs), or problems that may be caused by the distribution system (bacteria, lead, copper).
Whole-Body-Contact Recreation and Secondary Contact Recreation	Fecal Coliform or E. coli count	1-4	<u>Full</u> : Water Quality Standards not exceeded as a geometric mean for samples collected during seasons and flow conditions for which bacteria criteria apply. ⁹ <u>Non-Attainment</u> : Requirements for full attainment not met.
Irrigation, Livestock and Wildlife Water	Chemical	1-4	<u>Full</u> : Water Quality Standard not exceeded. ⁹ <u>Non-Attainment</u> : Requirements for full attainment not met

*Maximum Contaminant Level.

Other Types of Information

1. Observation and evaluation of waters for noncompliance with state narrative water quality criteria. These narrative criteria apply to both classified and unclassified waters and prohibit

Report Series No. 505, Sixteenth Report on the Joint FAO/WHO Expert Committee on Food Additives. Geneva 33 pp.

¹¹ Raw water samples used for compliance with source water criteria must be taken at a depth no deeper than the depth of the drinking water intake.

the following in waters of the state. Missouri's narrative water quality criteria as described in 10 CSR 20-7.031, Section (3) may be used to evaluate waters when a quantitative value can be applied to the pollutant.

- a. unsightly, putrescent or harmful bottom deposits
 - b. oil, scum and floating debris
 - c. unsightly turbidity or odor
 - d. substances causing toxicity to human, animal or aquatic life
 - e. human health hazard due to incidental contact
 - f. acute toxicity to livestock or wildlife when used as a drinking water supply
 - g. physical, chemical or hydrologic changes that impair the natural biological community
 - h. used tires, car bodies, appliances, demolition debris, used vehicles or equipment and any solid waste as defined by Missouri's Solid Waste Law
2. Habitat Assessments. Habitat assessment protocols for wadeable streams have been established and are made in conjunction with sampling of aquatic invertebrates and the analysis of aquatic invertebrates data. Missouri Department of Natural Resources will not use habitat assessment data alone for assessment purposes.

E. 303(d) Listing Considerations

1. Adding to the Existing List or Expanding the Scope of Impairment to a Previously Listed Water

New waters are added to the list following EPA guidelines in Section 1B and data evaluation rules in Table 1. Likewise, the listed portion of an impaired water can be increased based on recent monitoring data following the same guidelines. Likewise one or more new pollutants can be added to the listing for a water already on the list based on recent monitoring data following the same guidelines.

2. Deleting from the Existing List or Decreasing the Scope of Impairment to a Previously Listed Water.

Waters are deleted from the list following EPA guidelines in Section 1B and data evaluation rules in Table 1. Likewise, the listed portion of an impaired water can be decreased based on recent monitoring data following these same guidelines. Likewise, one or more pollutants can be deleted from the listing for a water already on the list based on recent monitoring data following these same guidelines.

3. Prioritization of Waters for TMDL Development

The department will prioritize development of TMDLs based on several variables including:

- severity of the water quality problem
- amount of time necessary to acquire sufficient data to develop the TMDL
- court orders, consent decrees or other formal agreements
- budgetary constraints
- amenability of the problem to treatment

The department's TMDL schedule will represent its prioritization. (ALRR p.63).

4. Resolution of Interstate/International Disagreements

The Missouri Department of Natural Resources will review the draft 303(d) lists of all other states with which we share border or other interstate waters (Missouri River, Mississippi River, Des Moines River and the St. Francis River). Where the listing in another state is different than in Missouri, the department will request the data upon which the listing in the other state is based. These data will be reviewed following all data evaluation guidelines previously discussed in this document and the Missouri list will or will not be changed pending the evaluation of this additional data.

Appendix A.

Excerpt from Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act. July 29, 2005. USEPA.pp.39-41.

G. How should statistical approaches be used in attainment determinations?

The state's methodology should provide a rationale for any statistical interpretation of data for the purpose of making an assessment determination.

1. Description of statistical methods to be employed in various circumstances

The methodology should provide a clear explanation of which analytic tools the state uses and under which circumstances. EPA recommends that the methodology explain issues such as the selection of key sample statistics (arithmetic mean concentration, median concentration, or a percentile), null and alternative hypotheses, confidence intervals, and Type I and Type II error thresholds. The choice of a statistic tool should be based on the known or expected distribution of the concentration of a pollutant in the segment (e.g. normal or log normal) in both time and space.

Past EPA guidance (1997 305b and 2000 CALM) recommended making non attainment decisions for “conventional pollutants” – TSS, pH, BOD, fecal coliform bacteria, and oil and grease – when more than “10% of measurements exceed the water quality criterion”. (However, EPA guidance has not encouraged use of the 10% rule with other pollutants, including toxics.) Use of this rule when addressing conventional¹² pollutants, is appropriate if its application is consistent with the manner in which the applicable WQC are expressed. An example of a WQC for which an assessment based on the ten percent rule would be appropriate is the EPA acute WQC for fecal coliform bacteria, applicable to protection of water contact recreational use. This 1976-issued WQC was expressed as, “...no more than ten percent of the samples exceeding 400 CFU per 100ml, during a 30-day period. Here, the assessment methodology is clearly reflective of the WQC.

¹² There are a variety of definitions for the term “conventional pollutant.” Wherever this term is referred to in this guidance, it means “a pollutant other than a toxic pollutant”.

On the other hand, use of the ten percent rule for interpreting water quality data is usually not consistent with WQC expressed either as: 1) instantaneous maxima not to be surpassed at any time, or 2) average concentrations over specified times. In the case of “instantaneous maxima (or minima) never to occur” criteria use of the ten percent rule typically leads to the belief that segment conditions are equal to or better than specified by the WQC, when they in fact are considerably worse. (That is, pollutant concentrations are above the criterion concentration a far greater proportion of the time than specified by the WQC). Conversely, use of this decision rule in concert with WQC expressed as average concentrations over specific times can lead to concluding that segment conditions are worse than WQC, when in fact they are not.

If the state applies different decision rules for different types of pollutants (e.g. toxic, conventional, and non-conventional pollutants) and types of standards (e.g. acute vs. chronic criteria for aquatic life or human health), the state should provide a reasonable rationale supporting the choice of a particular statistical approach to each of its different sets of pollutants and types of standards.

2. Elucidation of policy choices embedded in selection of particular statistical approaches and use of certain assumptions

EPA strongly encourages states to highlight policy decisions implicit in the statistical analysis that they have chosen to employ in various circumstances. For example, if hypothesis testing is used, the state should make its decision-making rules transparent by explaining why it chose either “meeting WQS” or “not meeting WQS” as the null hypothesis (rebuttable presumption) as a general rule for all waters, a category of waters, or an individual segment. Starting with the assumption that a water is “healthy” when employing hypothesis testing means that a segment will be identified as impaired, and placed in Category 4 or 5, only if substantial amounts of credible evidence exist to refute the presumption. By contrast, making the null hypothesis “WQS not being met” shifts the burden of proof to those who believe the segment is, in fact, meeting WQS.

Which ‘null hypothesis’ a state selects could likely create contrasting incentives regarding support for additional ambient monitoring among different stakeholders. If the null hypothesis is ‘meeting standards’, there were no previous data on the segment, and no additional existing and readily available data and information are collected, then the ‘null hypothesis’ cannot be rejected, and the segment would not be placed in Category 4 or 5. In this situation, those concerned about possible adverse consequences of having a segment declared ‘impaired’ might have little interest in collection of additional ambient data. Meanwhile, users of the segment would likely want to have the segment monitored, so they can be ensured that it is indeed capable of supporting the uses of concern. On the other hand, if the null hypothesis is changed to ‘segment not meeting WQS’, then those that would prefer that a particular segment not be labeled ‘impaired’ would probably want more data collected, in hopes of proving that the null hypothesis is not true.

Another key policy issue in hypothesis testing is what significance level to use in deciding whether to reject the null hypothesis. Picking a high level of significance for rejecting the null hypothesis means that great emphasis is being placed on

avoiding a Type I error (rejecting the null hypothesis, when in fact, the null hypothesis is true). This means that if a 0.10 significance level is chosen, the state wants to keep the chance of making a Type I error at or below ten percent. Hence, if the chosen null hypothesis is ‘segment meeting WQS’, the state is trying to keep the chance of saying a segment is impaired, when in reality it is not, under ten percent.

An additional policy issue is the Type II errors (not rejecting the null hypothesis, when it should have been). The probability of Type II errors depends on several factors. One key factor is the number of samples available. With a fixed number of samples, as the probability of Type I error decreases, the probability of a Type II error increases. States would ideally collect enough samples so the chances of making Type I and Type II errors are simultaneously small. Unfortunately, resources needed to collect such numbers of samples are quite often not available.

The final example of a policy issue that a state should describe is the rationale for concentrating limited resources to support data collection and statistical analysis in segments where there are documented water quality problems or where the combination of nonpoint source loadings and point source discharges would indicate a strong potential for a water quality problem to exist.

EPA recommends that, when picking the decision rules and statistical methods to be utilized when interpreting data and information, states attempt to minimize the chances of making either of the following two errors:

- Concluding the segment is impaired, when in fact it is not, and
- Deciding not to declare a segment impaired, when it is in fact impaired.

States should specify in their methodology what significance level they have chosen to use, in various circumstances. The methodology would best describe in ‘plain English’ the likelihood of deciding to list a segment that in reality is not impaired (Type I error if the null hypothesis is “segment not impaired”). Also, EPA encourages states to estimate, in their assessment databases, the probability of making a Type II error (not putting on the 303(d) list a segment that in fact fails to meet WQS), when: 1) commonly-available numbers of grab samples are available, and 2) the degree of variance in pollutant concentrations are at commonly encountered levels. For example, if an assessment is being performed with a WQC expressed as a 30-day average concentration of a certain pollutant, it would be useful to estimate the probability of a Type II error when the number of available samples over a 30 day period is equal to the average number of samples for that pollutant in segments state-wide, or in a given group of segments, assuming a degree of variance in levels of the pollutant often observed over typical 30 day periods.

Appendix B Statistical Considerations

The most recent EPA guidance on the use of statistics in the 303(d) Listing Methodology document is given in Appendix A. Within this guidance there are three major recommendations regarding statistics:

- Provide a description of which analytical tools the state uses under various circumstances

- When conducting hypothesis testing, explain the various circumstances under which the burden of proof is placed on proving the water is impaired and when it is placed on proving the water is unimpaired
- explain the level of statistical significance used under various circumstances

Description of Analytical Tools

The table below describes the analytical tools the department proposes and asks stakeholders to comment on. In general, the information in the right-hand three columns is negotiable and should encompass the area of discussion with stakeholders.

Table B-1. Description of Analytical Tools

Beneficial Use	Analyte	Decision Rule	Analytical Tool	Confidence Limit ¹³	Significance Level
Narrative Criteria	Color	Significantly greater than control stream	Hypothesis Test 2 Sample t test	not applicable	0.10
	Bottom Deposits	Solids of anthropogenic origin cover more than 10% of stream bottom	Hypothesis Test 1 Sided Confidence Limit	60% Lower Confidence Limit	0.40
Aquatic Life	Biological Monitoring	Significant alteration of composition or reduction in diversity	7 or fewer samples using DNR Invert. protocol: rate as unimpaired if 75% of scores are 16 or more, rate as impaired if 75% of scores are 14 or less, else rate as inconclusive.	not applicable	not applicable
			8 or more samples using DNR Invert. protocol: % of samples that score 16 or more is no less than 5% less than the percent for reference stream in that EDU.	not applicable	not applicable
			other biological monitoring” to be determined by type of data.		
	Toxic Chemicals	No more than one exceedence of acute or chronic criterion in 3 yrs	not applicable	not applicable	not applicable
Conventional chemicals		No more than 10% of samples exceed criterion	40 or fewer samples: binomial probability	90% LCL	0.10
			More than 40 samples: Estimated binomial probability, ‘z’ statistic	90% LCL	0.10

¹³ Where hypothesis testing is used, for datasets with five samples or fewer, a 75% confidence interval around the appropriate central tendencies will be used to determine use attainment status. Use attainment will be determined as follows: (1) If the criterion value is above this interval (all values within the interval are in conformance with the criterion), rate as unimpaired. (2) If the criterion value falls within this interval, rate as unimpaired and place in Category 3B. (3) If the criterion value is below this interval (all values within the interval are not in conformance with the criterion), rate as impaired.

Fish Consumpt.	Toxic Chemicals in water	Concentration does not exceed criterion	Hypothesis test 1 Sided Confidence limit	60% Upper Confidence Limit	0.40
	Toxic Chemicals in Tissue	Concentration in fillet samples does not exceed criterion			
			-3 or more samples: Hypothesis test 1 Sided Confidence Limit	60% Upper Confidence Limit	0.40
Drinking Water Supply (Raw)	Toxic Chemicals	Concentration does not exceed criterion	Hypothesis test 1 Sided Confidence limit	60% Upper Confidence Limit	0.40
Drinking Water Supply (Raw)	Non-toxic Chemicals	Concentration does not exceed criterion	Hypothesis test 1 Sided Confidence limit	60% Lower Confidence Limit	0.40
Drinking Water Supply (Finished)	Toxic Chemicals, Bacteria	Concentration does not exceed criterion	Methods stipulated by Safe Drinking Water Act		
Whole Body Contact and Secondary Contact Rec.	Bacteria	Antilog of log transformed data does not exceed criterion	Hypothesis test 1 Sided Confidence limit	60% Upper Confidence Limit	0.40
Irrigation & Livestock Water	Toxic Chemicals	Concentration does not exceed criterion	Hypothesis test 1 Sided Confidence limit	60% Lower Confidence Limit	0.40

Rationale for the Burden of Proof

Hypothesis testing is a common statistical practice. The procedure involves first stating a hypothesis you want to test, such as “the most frequently seen color on clothing at a St. Louis Cardinals game is red”, and then the opposite or null hypothesis, “red is not the most frequently seen color on clothing at a Cardinals game”. Then a statistical test is applied to the data (a sample of the predominant color of clothing worn by 200 fans at a Cardinals game on July 12) and based on an analysis of that data, one of the two hypotheses is chosen as correct.

In hypothesis testing, the burden of proof is always on the test hypothesis. In other words, there must be very convincing data to make us conclude that the null hypothesis is not true and that we must accept the test hypothesis. How convincing the data must be is stated as the “significance level” of the test. A significance level of 0.10 means that there must be at least a 90% probability that the test hypothesis is true before we can accept it and reject the null hypothesis.

When conducting hypothesis testing on water quality data, there are two general ways of stating the hypothesis. In the first, the test hypothesis is that the water is impaired and the null hypothesis is that water quality is unimpaired. In the second general case the test hypothesis is that the water is unimpaired and the null hypothesis is that the water is impaired. In the first case, the burden of proof requires convincing data that the water is impaired and in the second, convincing data that the water is unimpaired. Which of these two ways should be used?

The decision of how to state these hypotheses and assign the burden of proof should, in some way, relate to the consequences of making an incorrect decision based on a statistical test. When using a high level of significance like 0.05 or 0.01, the probability of accepting the test hypothesis when it was in fact, incorrect (Type I error) is very low. However, in this same test, the chance of accepting the null hypothesis when in fact, it was incorrect (Type II error) can be very high. If the data set was small, the Type II error could exceed 80-90%. Thus, particularly for small data sets, there can be a large discrepancy in error rates between Type I and Type II errors.

Thus in assigning the burden of proof, we should assign the lower error rate to the error that has the most serious consequences. The approach we are proposing in Table B-1 is, where hypothesis testing is used, to assign the burden of proof on showing that the water is not impaired where human health criteria are involved, and for other criteria assign the burden of proof on showing that the water is impaired.

As an example, atrazine data in Monroe City Route J Lake from 1997 to 2004 has the following data attributes: mean = 2.997 ug/L, standard deviation = 2.41, number of samples = 87. We have three options for statistical analysis of this data. One, make a direct comparison of the sample mean to the criterion value. Two, use a hypothesis test that places the burden of proof on showing the water is impaired or three, using a hypothesis test that places the burden of proof on showing the water is unimpaired. Using the sample mean indicates the lake is just meeting the 3.000 ug/l criterion value for atrazine and should be judged unimpaired, but we cannot assign a level of confidence to this decision. Calculating a lower confidence limit of the mean places the burden of proof on showing the water is impaired and the upper confidence limit on proving the water is unimpaired.

The formula for a 60% one-sided confidence limit is:

$$\text{upper confidence limit (UCL)} = \text{sample mean} + ((0.253)(\text{std.deviation})/\text{sq.root of sample size})$$
$$\text{lower confidence limit (LCL)} = \text{sample mean} - ((0.253)(\text{std.deviation})/\text{sq.root of sample size})$$

For Monroe City Lake the 60 % UCL is 3.062 ug/L and the LCL is 2.932 ug/L. Thus placing the burden of proof on showing the lake is impaired, we compare the LCL of 2.932 to the 3 ug/L criterion and conclude the lake is unimpaired by atrazine. The actual statistical statement is “there is a 60 % probability that the true mean is greater than or equal to 2.932 ug/L”. Placing the burden of proof on showing the lake is unimpaired, we compare the UCL of 3.062 to the 3.000 ug/L criterion and conclude the lake is impaired. The actual statistical statement is “there is a 60% probability that the true mean is less than or equal to 3.062 ug/L.”

Neither of these two statements is very reassuring that the lake is really meeting the atrazine standard since the mean value is so close to the criterion value. If the sample data is applied to the frequency distribution of the “z” statistic, the result shows that true mean has only a 50.4% chance of meeting the criterion. It seems unwise to judge a water to be in conformance with a human health criterion when there is only a 50.4% chance that this is true¹⁴.

For non-human health related criteria that are to be evaluated by hypothesis tests the burden of proof will be placed on showing that the water is impaired. The rationale for this decision is that the often significant social and or economic consequences that follow listing a water as impaired should not be imposed unless there is convincing evidence of impairment. Placing the burden of proof in this manner provides impetus to the department and other environmental agencies and organizations to focus monitoring efforts on these waters.

¹⁴ Monroe City Lake provides a good example of the value of an upper confidence limit. However, it should be noted that this data set is highly unusual in the similarity of the mean value (2.997) and the criterion value (3.000). In two other lakes in northeast Missouri with mean atrazine levels of 2.20 and 2.36, the UCL was less than the criterion value and use of the UCL, mean or LCL all lead to the conclusion that the water was unimpaired by atrazine. The message here is that we can use the UCL to judge impairment with the added confidence it provides that we are protecting human health, without greatly adding to the list of waters we judge to be impaired. This becomes increasingly true as sample size increases.

Level of Significance Used in Tests

The choice of significance levels is largely related to two concerns. The first, as with the wording of the test and null hypotheses, is concerned with matching error rates with the severity of the consequences of making a decision error. The second addresses the need to balance to the degree practicable, Type I and Type II error rates. For most circumstances, the department is recommending a 0.40 significance. We are not recommending a more stringent significance level such as 0.10 because of the increase in Type II error rate when going to this higher significance level.

For relatively small databases, the disparity between Type I and Type II errors can be large. The table below shows error rates calculated using the binomial distribution for two very similar situations. Type I error rates are based on a stream with a 10% exceedence rate of a standard and Type II error rates for a stream with a 15% exceedence rate of a standard. Note that choosing a Type I error rate of 0.05 rather than 0.10 increases an already very large Type II error rate by about ten percent. Also note that for a given Type I error rate, the Type II error rate declines as sample size increases.

Table B-2. Effects of Type I Error Rates and Sample Size on Type II Error Rates

No. of Samples	No. Meeting Standards	Type I Error Rate	Type II Error Rate	No. of Samples	No. Meeting Standards	Type I Error Rate	Type II Error Rate
6	5	.11	.78	4	3	.05	.89
11	9	.09	.78	9	7	.05	.86
18	15	.10	.72	15	12	.05	.82
25	21	.10	.68	21	17	.05	.80
				27	21	.05	.78

Other Statistical Considerations

Prior to calculation of confidence limits, the normality of the data set will be evaluated. If normality is improved by a data transformation, the confidence limits will be calculated on the transformed data.

Time of sample collection may be biased and interfere with an accurate measurement of frequency of exceedence of a criterion. Data sets composed mainly or entirely of storm water data, or data collected only during a season when water quality problems are expected could result in a biased estimate of the true exceedence frequency. In these cases, the department may use methods to estimate the true annual frequency and display these calculations whenever they result in a change in the impairment status of a water.

For waters judged to be impaired based on biological data where data evaluation procedures are not specifically noted in Table One, the statistical procedure used, test assumptions and results will be reported.

Appendix C

Examples of Statistical Procedures

Two Sample t Test for Color

Null Hypothesis: amount of color is no greater in test stream than in a control stream. (As stated, this is a one-sided test, meaning that we are only interested in determining whether or not the color level in the test stream is greater than in a control stream. If the null hypothesis had been “amount of color is different in the test and control streams” we would have been interested in determining if the amount of color was either less than or greater than the control stream, a two sided test).

Significance Level (also known as the alpha level): 0.10.

Data Set. Platinum-Cobalt color units data for the test stream and a control stream samples collected at each stream on same date.

Test Stream	70	45	35	45	60	60	80
Control Stream	50	40	20	40	30	40	75
Difference (T-C)	20	5	15	5	30	20	5

Statistics for the Difference: Mean = 14.28, standard deviation = 9.76, n = 7

Calculated “t” value = (square root of n)(mean)/standard deviation = 3.86

Tabular “t” value is taken from a table of the “t” distribution for 2 alpha (0.20) and n-1 degrees of freedom. Tabular “t” = 1.44.

Since calculated t value is greater than tabular t value, reject the null hypothesis and conclude that the test stream is impaired by color.

Statistical Procedure for Small Datasets for Mercury in Fish Tissue.

Example

Data Set: data in ug/Kg 150, 280, 130, 230, 450.

Mean = 248, standard deviation = 128.14, n=5.

Since the data set contains only five samples, a 75% confidence interval around the mean is calculated as: plus or minus $(1.15)(128.14)/\sqrt{5} = 65.9$

Thus the confidence interval is 182.1 to 313.9. Since the criterion value, 300 ug/Kg, falls within the 75% confidence interval, this waterbody would be placed in Category 3B.